AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 112:149–158 (2000)

# Origin of YAP+ Lineages of the Human Y-Chromosome

CLAUDIO M. BRAVI, GRACIELA BAILLIET,
VERÓNICA L. MARTINEZ-MARIGNAC, AND NÉSTOR O. BIANCHI*
*Multidisciplinary Institute of Cell Biology (IMBICE),*
*1900 La Plata, Argentina*

*KEY WORDS* YAP+ chromosomes; YAP+ lineages; Y-specific *Alu* inserts

*ABSTRACT* We screened a total of 841 Y-chromosomes representing 36 human populations of wide geographical distribution for the presence of a Y-specific *Alu* insert (YAP+ chromosomes). The *Alu* element was found in 77 cases. We tested 5 biallelic and 8 polyallelic markers in 70 out of the 77 YAP+ chromosomes. We could identify the existence of a hierarchical and chronological structuring of ancestral and derived YAP+ lineages, giving rise to 4 haplogroups, 14 subhaplogroups and 60 haplotypes. Moreover, we propose a monophyletic origin for each one of the YAP+ lineages. Out-of-Africa and out-of-Asia models have been suggested to explain the origin and evolution of ancestral and derived YAP+ elements. We analyze the evidence supporting these two hypotheses, and we conclude that the information available does not allow one to decide between the out-of-Asia or out-of-Africa models. Am J Phys Anthropol 112:149–158, 2000. 

The Y-specific region of Y-chromosomes is haploid and patrilineally transmitted along generations. As this region does not undergo recombination, all Y-specific genes and markers are in linkage disequilibrium, with mutations being the only potential source of variation between the male ancestor and his male offspring.

Several *Alu* inserts known to occur in humans after the human/ape divergence have been extensively used to infer the evolution of modern human populations (Batzer et al., 1994; Stoneking et al., 1997; Hammer, 1994, 1995; Whitfield et al., 1995). Among these human-specific elements, one of them is particularly interesting due to its location in the nonrecombinant region (DYS287 locus) of the human Y-chromosome; YAP+ and YAP− acronyms identify the presence or absence (ancestral state) of the insert respectively (Hammer, 1994; Spurdle et al., 1994).

YAP+ chromosomes appear at low frequency (<10%) in some Asian, Oceanian, and Amerindian populations, at intermediate frequency (11–30%) in Caucasians and Japanese, and at high frequency (>31%) in Tibetans and several African populations (Altheide and Hammer, 1997; Hammer, 1994; Hammer et al., 1998; Spurdle et al., 1994; Bianchi et al., 1997; Karafet et al., 1997; Thomas et al., 1998). Since its origin, this *Alu* domain accumulated several biallelic additional mutations: two C > T transitions at 338 bp (PN1) and 1,682 bp (PN2) loci, and a deletion of the Poly-A tail (S for short tail). Finally, outside the *Alu* domain, two other mutations, a G > A transition at 4,064 bp of the SRY domain (Whitfield et al., 1995) and an A > G transition at the 168 bp position of the DYS271 locus (Seielstad et al., 1994) were also informative (Hammer,

Grant sponsor: National Council of Research, Argentina; Grant sponsor: Research Commission of the province of Buenos Aires, Argentina; Grant sponsor: Argentine-Brazilian Center for Biotechnology; Grant sponsor: National Agency for the Development of Science and Technology, Argentina.

*Correspondence to: Dr. Néstor O.Bianchi, IMBICE, Calle 526 e/10 y 11, 1900 La Plata, Argentina.
E-mail: imbice@satlink.com or bianchi@satlink.com

Received 4 January 1999; accepted 24 January 2000.

1995; Altheide and Hammer, 1997; Bravi et al., 1997a,b).

Besides the above biallelic loci, we analyze here 8 additional polyallelic Y-specific markers (one alphoid and seven microsatellite variants) in order to obtain a further insight on the origin and evolution of YAP+ lineages. The use of these markers allows us to identify YAP+ haplogroups, subhaplogroups, and haplotypes to establish the monophyletic origin of the lineages and to assess the hypotheses regarding the geographic origin and age of the ancestral YAP+ chromosome.

## MATERIALS AND METHODS

We studied a total of 841 Y-chromosomes representing different populations and geographic regions. Details of the populations analyzed are given in Table 1. Allele DYS199*T is specific to Native Americans (Underhill et al., 1996; Bianchi et al., 1998); 64% of the Amerindian Y-chromosomes in our series showed this marker. In this report, all cases of YAP+ chromosomes in native Americans had the ancestral allele DYS199*C, indicating that these chromosomes are the result of European (haplogroup C) or African (haplogroup D) admixture.

Samples were provided by M. Hammer (1 Tibetan and 1 Japanese, 1 Bantu and 2 Gambians), Y.-F. Chris Lau (20 Chinese and all Laotians, Cambodians, Thai, Vietnamese, and Philippines), M. Sans (Afro-Uruguayans), G. Cantos (Afro-Ecuatorians), R. Herrera (Afro-USA, Pakistanis and Bengalis, Chimila, Maya, Zuñi, Sioux, and Navajo), J. Ferrer (Ayoreo and Lengua), P. Zukas (Mocovies), F. Rothhammer (Huilliche and Pehuenche), F. Carnese (Mapuche, Tehuelche, Wichi, Chorote, and Toba), Nippon University Association of La Plata (12 Japanese), Lebanese Society of La Plata and Argentine Arab Home of Berisso (Lebanese), Syrian-Orthodox Association of La Plata (Syrians), Centre d'Etude du Polymorphisme Humain (CEPH pedigrees), Coriell Cell Repositories (Pygmies and Melanesian), and the IMBICE DNA repository (La Plata and Jews). Coriell samples and samples donated by M. Hammer were also included in series previously reported by Hammer et al. (1997, 1998). As far as we know, CEPH samples were not tested previously for YAP+ chromosomes. Therefore, 829 out of the 841 Y-chromosomes tested for *Alu* inserts and the information on the αh system and microsatellites are new data produced by our group.

*TABLE 1. Origin of individuals analyzed for YAP+ chromosomes*

| | Populations analyzed | | YAP+ |
|---|---|---|---|
| East Asians | Tibetan | 1 | 1 |
| | Japanese | 13 | 4 |
| | Chinese | 23 | 0 |
| | Laotian | 7 | 0 |
| | Cambodian | 3 | 0 |
| | Taiwanese | 1 | 0 |
| South East Asians | Thai | 1 | 0 |
| | Philippine | 12 | 0 |
| | Vietnamese | 3 | 0 |
| South Asians | Bengali | 61 | 0 |
| | Pakistani | 91 | 0 |
| West Asians | Jew | 18 | 4 |
| | Syrian | 2 | 0 |
| | Lebanese | 6 | 0 |
| Oceanians | Melanesian | 2 | 0 |
| Africans | Bantu | 1 | 1 |
| | Gambian | 2 | 2 |
| | Pygmy | 5 | 2 |
| | Afro-Uruguayan | 17 | 7 |
| | Afro-Ecuatorean | 24 | 9 |
| | Afro-USA | 24 | 16 |
| Europeans | CEPH | 68 | 3 |
| South Americans | La Plata (admixed) | 199 | 22 |
| Native Americans | Huilliche | 20 | 0 |
| | Pehuenche | 21 | 1 |
| | Mapuche | 27 | 2 |
| | Tehuelche | 16 | 0 |
| | Mocovi | 26 | 0 |
| | Wichi | 25 | 1 |
| | Chorote | 11 | 0 |
| | Toba | 6 | 1 |
| | Ayoreo | 10 | 0 |
| | Lengua | 30 | 0 |
| | Chimila | 12 | 1 |
| | Maya | 12 | 1 |
| | Sioux | 16 | 0 |
| | Zuñi | 16 | 0 |
| | Navajo | 10 | 0 |
| Total | | 841 | 77 |

YAP+ and YAP− chromosomes were identified according to Hammer and Horai (1995). All YAP+ cases were tested for αh variants (27 forms) (Santos et al., 1996), for SRY 4,064 (Whitfield et al., 1995), Poly-A tail, PN1, PN2 (Hammer, 1995), and DYS271 (Seielstad et al., 1994) polymorphisms, and for the allelic form of the following microsatellites: DYS19 (tetranucleotide, 10 alleles), DYS389a (tetranucleotide, 7 alleles), DYS389b (tetranucleotide, 9 alleles), DYS390 (tetranucleotide, 10 alleles),

DYS391 (tetranucleotide, 6 alleles), DYS392 (tetranucleotide, 8 alleles), and DYS393 (trinucleotide, 6 alleles) (Kayser et al., 1997; de Kniff et al., 1997).

Alphoid (αh), DYS271 polymorphisms, and microsatellite allelic forms were tested as indicated in Santos et al. (1996), Seielstad et al. (1994), Kayser et al. (1997), and de Kniff et al. (1997). Microsatellite alleles were identified against sequenced standards kindly provided by Lutz Roewer and Manfred Kayser (Institute of Legal Medicine, Humboldt University, Berlin, Germany). The A > G transition at 4,064 bp of the SRY domain (Whitfield et al., 1995) was determined by PCR, using the pair of primers SRY 3,920 (forward) 5′-AGCACA TTAGCTGGTATGACA -3′ and SRY 4,425 (reverse) 5′-CTCTTTATGGCAAGACTTACG -3′, with 55°C annealing temperature. After digestion of the PCR fragment with *Bsr*BI, the ancestral allele (A) produces two restriction fragments of 359 bp and 146 bp, while the derived allele generates a single fragment of 505 bp due to lack of restriction. The PN1 C > T transition was detected by PCR followed by restriction of the amplified fragment. Primers used were PN1F (forward) 5′-CATAATTTCATTTTCCCTAT TGC-3′, and PN1B (reverse) 5′-GTCCTCTCCTTATTAACGTAA-3′, with 55°C annealing temperature. After digestion with *Alu*I, the 218-bp amplicon results in two fragments of 145 bp and 73 bp if the derived T allele is present. The PN2 C > T transition was detected by PCR and restriction, using a mismatched primer in order to produce a new *Hae*III site exhibiting the polymorphism in the last C of the GGCC recognition site. The change of this C into T inhibits *Hae*III digestion. Primers used were PN2F (forward) 5′-AGGAGCATTAATAAAACTAAgGC-3′ (g indicates the mismatch position) and PN2B (reverse) 5′-CTTCACTACCAGCCTAAGTAC-3′, with an annealing temperature of 56°C. Restriction in the ancestral allele produces 156-bp and 22-bp fragments; lack of restriction due to the C > T transition generates a 178-bp fragment. The test for the presence of a long (L) or short (S) Poly-A tail was made by *Hae*III restriction of the PCR fragment used for diagnosis of YAP+ chromosomes. In the presence of an L tail, the *Hae*III fragments produced have 213 bp, 80 bp, 50 bp, and 130 bp; the size of the 130-bp fragment decreases to 110 bp if the tail is S. L and S alleles in this report correspond to the L (46 Poly-A) and S (26 Poly-A) of Hammer et al. (1997).

Microsatellite alleles in tables and text are identified by the number of repeats; the αh variants are identified with Roman numerals (Kayser et al., 1997; Santos et al., 1996). Nomenclature for the YAP+ haplogroups follows that of Altheide and Hammer (1997); subhaplogroups are identified by Arabic numerals in the form of subscripts, and haplotypes by Arabic numerals.

## RESULTS

We found 77 YAP+ chromosomes out of the 841 samples analyzed. Table 1 illustrates the geographical origin of these 77 cases. Japanese and populations of African ancestry were the ones showing the highest frequencies of YAP+ chromosomes. In 70 YAP+ chromosomes we could make the full testing of the 13 polymorphic markers employed in this report; these are the cases included in Table 1.

Genotyping of the SRY 4,064, Poly-A, PN2, DYS271, and PN1 polymorphisms allowed us to define the four haplogroups 3G, 3A, 4, and 5 according to Altheide and Hammer (1997). The ancestral haplogroup 3G was characterized by the presence of ancestral alleles in all the above loci. Five Y-chromosomes belonged to this group (Table 2). A G > A transition in the SRY 4,064 locus identifies the haplogroup 3A, which comprises 7 cases (Table 2). The 31 YAP+ cases in haplogroup 4 exhibited an S Poly-A and the mutated allele PN2*T (Table 2). Finally, the derived alleles DYS271*G and PN1*T identify the 27 YAP+ chromosomes in haplogroup 5 (Table 2). No cases with medium (M, 36 bp) or very short (VS, 19 bp) Poly-A tails (Hammer et al., 1997) were detected in our series.

The αh system served to identify subhaplogroups within each haplogroup. We reported elsewhere that αhV and αII are the only two forms present in both YAP+ and YAP− chromosomes, and that αhV is the ancestral state (Santos et al., 1996). Thus, we considered 3G1 (αhV) and 3G2 (αhII) as the ancestral and derived haplogroups, re-

*TABLE 2. YAP+ haplogroups, subhaplogroups and haplotypes*

| Haplog | SRY 4,064 | Poly-A | PN 2 | DYS271 | PN 1 | Sub-hap | αh | 19 | 389a | 389b | 390 | 391 | 392 | 393 | Frequency | Haplotype | Origin |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3G | G | L | C | A | C | $3G_1$ | V | 15 | 10 | 27 | 24 | 10 | 11 | 12 | 1 | 1 | Tibetan |
| | G | L | C | A | C | $3G_2$ | II | 16 | 10 | 27 | 25 | 10 | 11 | 13 | 1 | 2 | Japanese |
| | G | L | C | A | C | $3G_2$ | II | 17 | 10 | 27 | 26 | 10 | 11 | 13 | 1 | 3 | Japanese |
| | G | L | C | A | C | $3G_2$ | II | 17 | 11 | 28 | 25 | 10 | 11 | 13 | 1 | 4 | Japanese |
| | G | L | C | A | C | $3G_2$ | II | 15 | 10 | 26 | 25 | 10 | 11 | 13 | 1 | 5 | Japanese |
| 3A | A | L | C | A | C | $3A_1$ | V | 14 | 9 | 25 | 25 | 10 | 11 | 13 | 1 | 6 | Pygmy |
| | A | L | C | A | C | $3A_1$ | V | 14 | 9 | 26 | 25 | 10 | 11 | 13 | 1 | 7 | Bantu |
| | A | L | C | A | C | $3A_1$ | V | 14 | 9 | 25 | 24 | 10 | 11 | 13 | 1 | 8 | Afro-USA |
| | A | L | C | A | C | $3A_1$ | V | 14 | 9 | 25 | 25 | 12 | 11 | 13 | 1 | 9 | Afro-USA |
| | A | L | C | A | C | $3A_1$ | V | 16 | 9 | 27 | 23 | 9 | 12 | 13 | 1 | 10 | La Plata |
| | A | L | C | A | C | $3A_1$ | V | 14 | 11 | 29 | 25 | 10 | 11 | 13 | 1 | 11 | Gambia |
| | A | L | C | A | C | $3A_1$ | V | 15 | 8 | 27 | 23 | 10 | 11 | 13 | 1 | 12 | Afro-USA |
| 4 | A | S | T | A | C | $4_1$ | V | 13 | 10 | 27 | 24 | 10 | 11 | 13 | 2 | 13 | La Plata (1), La Plata Jew (1) |
| | A | S | T | A | C | $4_1$ | V | 13 | 10 | 27 | 23 | 10 | 11 | 13 | 2 | 14 | La Plata (1), La Plata Jew (1) |
| | A | S | T | A | C | $4_1$ | V | 13 | 10 | 28 | 24 | 9 | 11 | 13 | 1 | 15 | La Plata |
| | A | S | T | A | C | $4_1$ | V | 13 | 9 | 27 | 23 | 10 | 11 | 13 | 1 | 16 | Afro-USA |
| | A | S | T | A | C | $4_1$ | V | 13 | 10 | 27 | 25 | 9 | 11 | 14 | 1 | 17 | La Plata Jew |
| | A | S | T | A | C | $4_1$ | V | 14 | 10 | 27 | 24 | 11 | 12 | 13 | 1 | 18 | Pehuenche |
| | A | S | T | A | C | $4_1$ | V | 14 | 10 | 29 | 24 | 10 | 11 | 13 | 1 | 19 | Mapuche |
| | A | S | T | A | C | $4_1$ | V | 13 | 11 | 28 | 25 | 10 | 11 | 13 | 1 | 20 | La Plata |
| | A | S | T | A | C | $4_1$ | V | 13 | 10 | 28 | 24 | 12 | 11 | 13 | 1 | 21 | La Plata |
| | A | S | T | A | C | $4_1$ | V | 14 | 9 | 26 | 24 | 10 | 11 | 12 | 1 | 22 | CEPH #66 |
| | A | S | T | A | C | $4_1$ | V | 14 | 10 | 28 | 25 | 9 | 12 | 13 | 1 | 23 | La Plata |
| | A | S | T | A | C | $4_1$ | V | 15 | 9 | 26 | 24 | 10 | 11 | 12 | 1 | 24 | Afro-Uruguay |
| | A | S | T | A | C | $4_1$ | V | 13 | 11 | 28 | 25 | 9 | 11 | 14 | 1 | 25 | La Plata Jew |
| | A | S | T | A | C | $4_2$ | III | 13 | 11 | 28 | 24 | 9 | 12 | 13 | 1 | 26 | La Plata |
| | A | S | T | A | C | $4_3$ | II | 13 | 11 | 26 | 24 | 9 | 11 | 13 | 1 | 27 | La Plata |
| | A | S | T | A | C | $4_3$ | II | 13 | 11 | 27 | 24 | 9 | 11 | 13 | 1 | 28 | La Plata |
| | A | S | T | A | C | $4_4$ | XII | 13 | 10 | 27 | 24 | 10 | 11 | 13 | 4 | 29 | La Plata (3), Wichi (1) |
| | A | S | T | A | C | $4_4$ | XII | 13 | 10 | 27 | 23 | 10 | 11 | 13 | 1 | 30 | La Plata |
| | A | S | T | A | C | $4_4$ | XII | 13 | 11 | 28 | 24 | 10 | 11 | 13 | 1 | 31 | La Plata |
| | A | S | T | A | C | $4_4$ | XII | 13 | 9 | 26 | 24 | 10 | 11 | 13 | 1 | 32 | La Plata |
| | A | S | T | A | C | $4_4$ | XII | 13 | 10 | 27 | 23 | 11 | 11 | 13 | 1 | 33 | La Plata |
| | A | S | T | A | C | $4_4$ | XII | 14 | 11 | 27 | 25 | 10 | 11 | 14 | 1 | 34 | CEPH #1423 |
| | A | S | T | A | C | $4_5$ | XVI | 13 | 11 | 27 | 23 | 9 | 11 | 13 | 1 | 35 | La Plata |
| | A | S | T | A | C | $4_5$ | XVI | 13 | 11 | 27 | 25 | 9 | 11 | 13 | 1 | 36 | La Plata |
| | A | S | T | A | C | $4_5$ | XVI | 13 | 11 | 26 | 24 | 9 | 11 | 13 | 1 | 37 | La Plata |
| | A | S | T | A | C | $4_5$ | XVI | 13 | 10 | 26 | 24 | 10 | 11 | 13 | 1 | 38 | Mapuche |
| 5 | A | S | T | G | T | $5_1$ | IX | 15 | 10 | NULL | 21 | 10 | 11 | 13 | 1 | 39 | Afro-Uruguay |
| | A | S | T | G | T | $5_1$ | IX | 15 | 11 | 28 | 21 | 10 | 11 | 14 | 1 | 40 | Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 15 | 10 | 28 | 21 | 10 | 11 | 13 | 1 | 41 | Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 15 | 10 | 27 | 21 | 10 | 11 | 14 | 1 | 42 | Chimila |
| | A | S | T | G | T | $5_1$ | IX | 16 | 11 | 28 | 21 | 10 | 11 | 14 | 5 | 43 | Afro-Ecuador |
| | A | S | T | G | T | $5_1$ | IX | 16 | 9 | 28 | 21 | 10 | 11 | 14 | 1 | 44 | Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 15 | 11 | 29 | 21 | 10 | 11 | 14 | 1 | 45 | Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 17 | 10 | 28 | 21 | 10 | 11 | 14 | 2 | 46 | Afro-Ecuador, Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 15 | 10 | 28 | 21 | 11 | 11 | 13 | 1 | 47 | Afro-Ecuador |
| | A | S | T | G | T | $5_1$ | IX | 15 | 10 | 27 | 22 | 10 | 11 | 13 | 1 | 48 | Afro-Ecuador |
| | A | S | T | G | T | $5_1$ | IX | 16 | 11 | 28 | 22 | 10 | 11 | 14 | 1 | 49 | Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 17 | 10 | NULL | 20 | 10 | 11 | 14 | 1 | 50 | Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 17 | 10 | 27 | 21 | 10 | 11 | 14 | 1 | 51 | Afro-USA |
| | A | S | T | G | T | $5_1$ | IX | 15 | 10 | 27 | 21 | 10 | 13 | 15 | 1 | 52 | Afro-Uruguay |
| | A | S | T | G | T | $5_1$ | IX | 15 | 10 | 26 | 21 | 10 | 12 | 15 | 1 | 53 | CEPH #106 |
| | A | S | T | G | T | $5_1$ | IX | 18 | 10 | 28 | 21 | 10 | 11 | 12 | 1 | 54 | Afro-Uruguay |
| | A | S | T | G | T | $5_1$ | IX | 17 | 10 | 27 | 21 | 10 | 11 | 14 | 1 | 55 | Afro-USA |
| | A | S | T | G | T | $5_2$ | XXIII | 14 | 10 | 27 | 21 | 10 | 11 | 15 | 1 | 56 | Toba |
| | A | S | T | G | T | $5_3$ | XXVI | 15 | 10 | 27 | 21 | 10 | 11 | 14 | 1 | 57 | Pygmy |
| | A | S | T | G | T | $5_4$ | XXVII | 15 | 10 | 27 | 21 | 10 | 11 | 14 | 1 | 58 | Maya |
| | A | S | T | G | T | $5_5$ | XIV | 15 | 10 | 28 | 21 | 10 | 11 | 13 | 1 | 59 | Afro-Ecuador |
| | A | S | T | G | T | $5_6$ | II | 15 | 10 | 28 | 21 | 10 | 11 | 13 | 1 | 60 | Afro-USA |
| Total | | | | | | | | | | | | | | | 70 | | |

spectively (Table 2, Fig. 1). In haplogroup 3A, all cases were αhV; thus, subhaplogroup 3A1 was the only one detected in this set of YAP+ chromosomes (Table 2, Fig.1). Alphoids αhII, III, XII, and XVI derive from αhV through one (form III) or two (form II) deletions, or via one mutation (forms XII and XVI) (Santos et al., 1996). Therefore, 41 (αhV) is ancestral, and 42–45 are the derived subhaplogroups (Table 2, Fig.1).

Alphoid forms αhXXIII, XXVI, and XXVII are known to derive from αhIX through deletions (αhXXIII) or point mutations (αhXXVI–XXVII) (Santos et al., 1996; Bailliet, unpublished findings). Moreover, αhIX is known to be a derivative form of αhV (Santos et al., 1996). Thus, two different alternatives may explain the subhaplogroup pattern in haplogroup 5: 1) αhV is the ancestral form; in this case, the lack of αhV chromosomes in this haplogroup is due to lineage loss or to insufficient sampling; or 2) the two mutations defining group 5 (DSY271*G and PN1*T) occurred in an αhIX chromosome, which is the ancestral form in this haplogroup. In the first alternative, the form αhII derives from αhV via two deletions, while in the second alternative αhII is assumed to be a derivative of αhXIV via two deletions in the same two loci involved in the V > II change (Santos et al., 1996). Due to the high frequency of αhIX chromosomes, we tentatively assume this form to be the ancestral one (51 subhaplogroup) (Table 2, Fig. 1).

The allele association for the 7 microsatellite loci analyzed allowed us to identify the haplotypes within each subhaplogroup (Table 2, Fig. 1). Since several individuals shared the same haplotypes (haplotypes 13, 14, 46, 29, and 43 were shared by 2, 2, 2, 4, and 5 males, respectively), we could only identify 60 different haplotypes out of the 70 Y-chromosomes analyzed.

## DISCUSSION

### Monophyletic origin of haplo- and subhaplogroups

The possibility of recurrence for *Alu* markers is almost null. Therefore, all YAP+ chromosomes are known to derive from a single individual in whom the insertion took place for the first time (Hammer, 1995).

Point mutations and deletions may be recurrent, but the frequency of these phenomena is usually low enough as to make it certain that most, if not all, the individuals in haplogroups 3A, 4, and 5 derive also from the first individual who underwent the haplogroup-specific mutation. Haplogroups 4 and 5 are peculiar in the sense that they are defined by the coexistence of two mutations: S and PN2 in the branching from 3A to 4, and DYS271 plus PN1 in the branching from 4 to 5 (Table 2, Fig. 1). In these cases it is clear that haplogroups 4 and 5 derived from the individual who had the second mutation in each pair of haplogroup-specific mutations.

If we add the new YAP+ cases in this paper to the YAP+ cases reported by Hammer et al. (1997), we have a total of 152 YAP+ chromosomes belonging to haplogroup 4 and 245 YAP+ chromosomes belonging to haplogroup 5. These sample sizes are sufficient to ensure the detection of one of the two mutated alleles at a frequency of 0.0046 with 95% and 99% confidence for haplogroups 4 and 5, respectively (Chakraborty, 1992). Therefore, the presence of individuals showing only one mutation of the pair, in case they exist, should occur at frequencies lower than 4.6‰. Thus, there is a strong possibility that the extinction of the older lineage may be the cause explaining the apparent concurrence of two mutations in the branching of 3A to 4 and 4 to 5.

A direct assessment of the mutation rate of the alphoid system shows that the change of αh forms occurs at a rate lower than $1.6 \times 10^{-4}$ (Bianchi et al., 1998). Thus, most or all subhaplogroup lineages probably derive from single individuals who for the first time showed the change of one αh form to another.

The monophyletic origin of YAP+ lineages is in good agreement with predictions formulated by von Haeseler et al. (1995) and Hardpending et al. (1998). In this regard we can envisage the evolution of YAP+ chromosomes as a chronological and hierarchical succession of severe bottlenecks followed by expansions. In a process of this sort, each lineage generating a haplogroup, subhaplogroup, or haplotype originates in a single individual and then grows through expan-

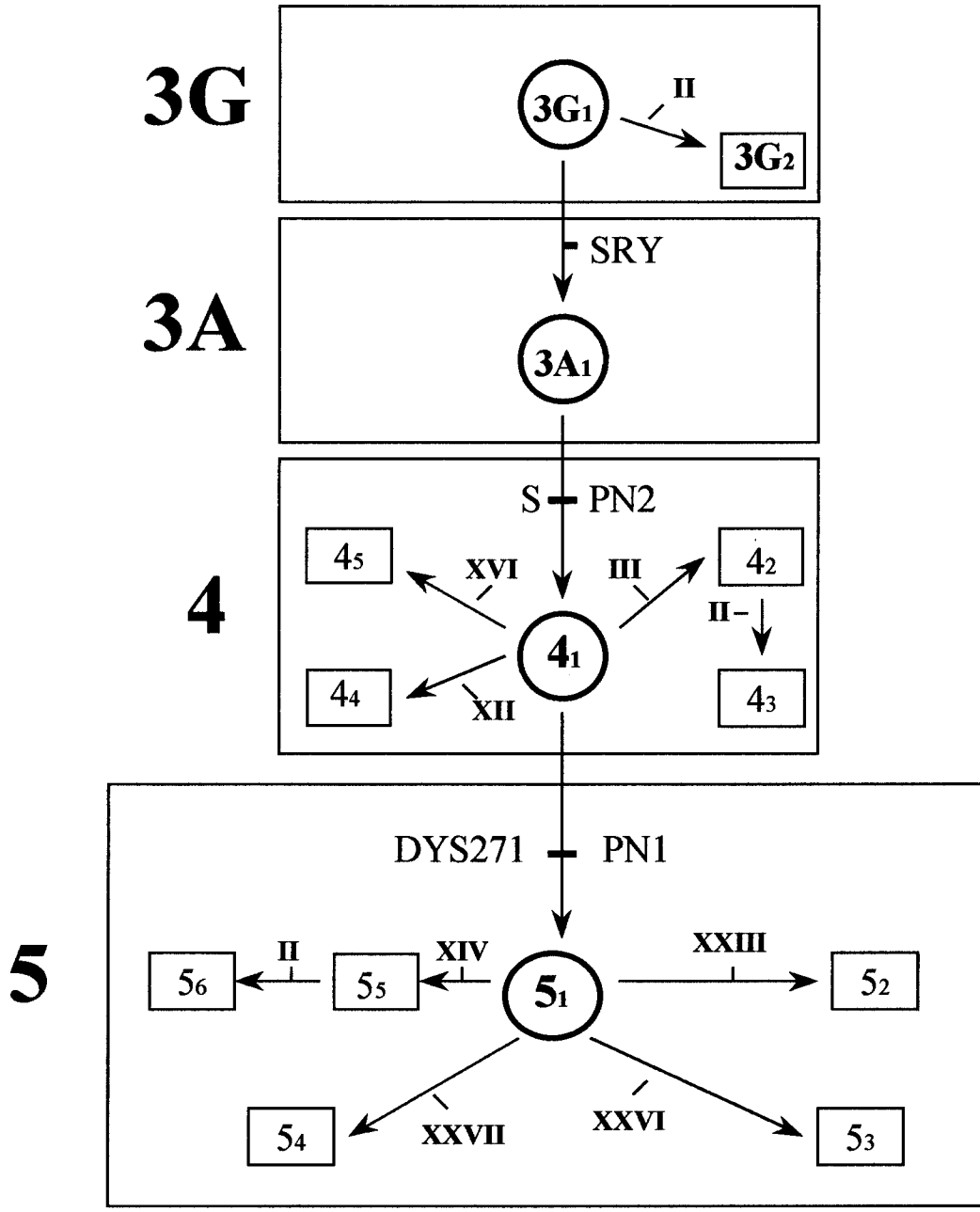


Fig. 1. Evolutionary pathway of YAP+ chromosomes. 3G, 3A, 4, and 5 indicate haplogroups. $3G_1$, $3A_1$, $4_1$–$4_5$, and $5_1$–$5_6$ indicate subhaplogroups. Arrows connect ancestral and derived forms. Acronyms in arrow shafts indicate mutations producing the derived forms. Two mutations connected by a bar indicate impossibility of detecting the chronology of appearance. SRY 4,064, G > A transition; S, Poly-A tail deletion; PN2, C > T transition; DYS271, A > G transition; PN1, C > T transition; Roman numerals correspond to αh forms (see Table 2). Haplotypes included in each subhaplogroup are detailed in Table 2.

sion, giving rise in some cases to the coexistence of the ancestral and derived lineages, to the partial or total replacement of the ancestral lineage by the derived one, or, in some occasions, to the total loss of the ancestral and derived forms. The finding of haplogroups 3A, 4, and 5 in Africa or subhaplogroups $3G_{1-2}$ in Asia, $4_1$–$4_5$ in Europe, and $5_1$–$5_6$ in Africa are examples of the coexistence of ancestral and derived forms. Lack of 3A haplogroups in Asia or 3G haplogroups in Africa (see below) suggests total replacement of the ancestral by the derived form. S, PN2, DYS271, and PN1 variants indicate a chronological succession of four mutational steps. Yet our series, as well as that of Altheide and Hammer (1997), only shows cases with two (S/PN2) or four (S/PN2/DYS271/PN1) mutations, implying the total or almost total replacement of cases with one mutation by the derived forms with two mutations, and the replacement of Y-chromosomes with three mutations by derived chromosomes with four mutations.

### Geographical origin and age of YAP+ chromosomes

In regard to the geographical distribution of ancestral and derived YAP+ chromosomes, our data are in agreement with those of Hammer et al. (1997, 1998) and Altheide and Hammer (1997). Haplogroup 3G was restricted to Tibetans and Japanese, and haplogroup 3A was only found in Africans. In our series, haplogroup 4 was prevalent for Europeans, while in Altheide and Hammer (1997), it was almost equally distributed between Europeans and Africans; moreover, this haplogroup has also been reported in Middle Easterners (Hammer et al., 1997; Thomas et al., 1998). Finally, the most recent haplogroup 5 was mainly detected in individuals of African ancestry (Table 2).

Lately, Altheide and Hammer (1997) proposed three models explaining the origin and evolution of YAP+ chromosomes. These models take into account the fact that the most ancestral haplogroup, 3G, is only found in Japanese and Tibetans. In the Asian founder model, it is assumed that haplogroup 3G originated in Asia and that haplogroup 3A appeared in the deme migrating into Africa, with subsequent loss of haplogroup 3G during migration or after entering Africa. In the Asian/Asian model, haplogroups 3G and 3A originated in Asia; after the migration of a 3A deme into Africa, the form 3A became extinct in the Asian area of origin. Finally, in the African model, haplogroup 3G emerged in Africa, to become lost after migration to Asia (Altheide and Hammer, 1997; Hammer et al., 1998).

Figure 2 illustrates the three models proposed for the geographical origin of ancestral and derived YAP+ haplogroups.

Models I, II, and III require the same number of migrations, haplogroup losses, and mutations. Therefore, it is difficult to decide which of these models is applicable to the early evolutionary stages of YAP+ elements.

Altheide and Hammer (1997) favor Asia as the site of origin of the Y-specific *Alu* insertion, based on the detection of haplogroup 3G in Asia and not in other continents, and based on the finding of a higher genetic diversity in haplogroup 3G when compared with the diversity of haplogroup 3A. Since haplogroup 3A is a derivative form of 3G, the increased diversity of 3G represents an additional demonstration of ancestrality but not of geographic origin. Thus, so far, the hypothesis of the Asian origin of YAP+ chromosomes rests only on the specific Asian location of haplogroup 3G. Yet, the lack of ancestral YAP+ chromosomes in regions between Japan and Tibet, and Tibet and Africa, is rather puzzling and makes it difficult to decide in favor of the out-of-Asia or out-of-Africa hypotheses.

The estimated age of the ancestral YAP+ chromosome may perhaps provide additional evidence on the geographical origin of the Y-specific *Alu* insertion. Hammer (1995) used five Y-specific biallelic markers to construct the evolutionary tree of YAP haplotypes. In that report, Hammer (1995) proposed Africa as the continent of origin of the ancestral Y and the ancestral YAP+ chromosomes, estimated the age of the ancestral Y-chromosome at 188,000 years (95% confidence interval of 411,000–51,000 years), and suggested 141,000 years ($\pm$81,000 years) as the age of YAP+ chromosomes; these dates

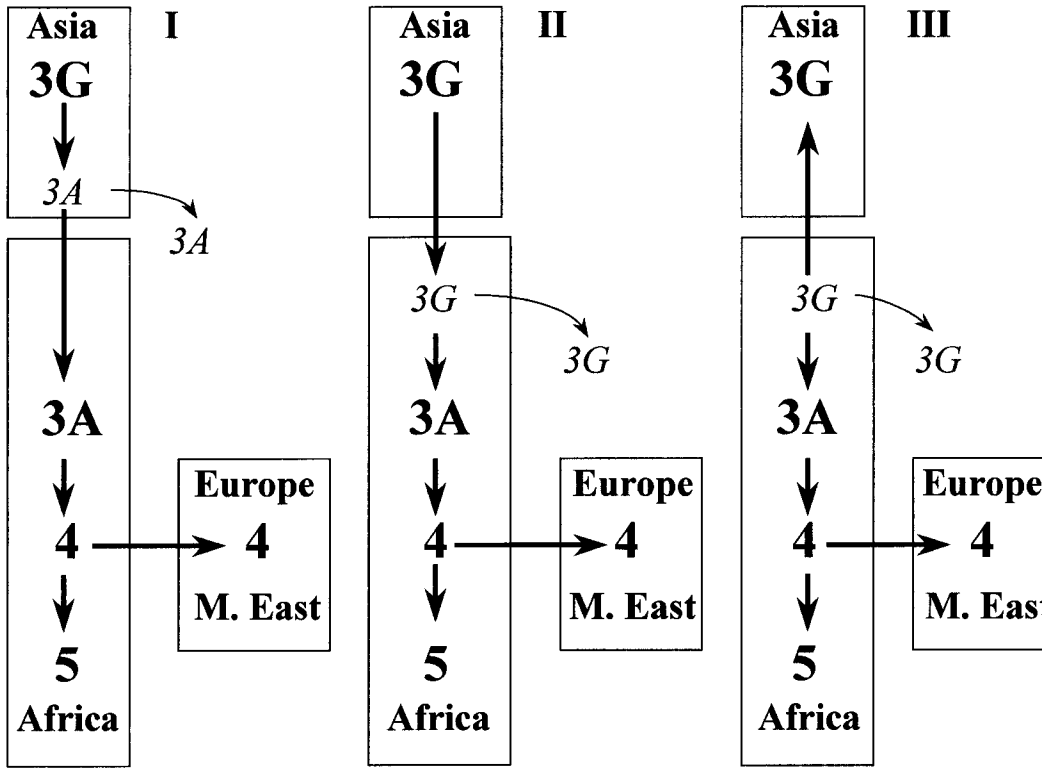


Fig. 2. Models regarding the geographical origin and evolution of YAP+ chromosomes. 3G, 3A, 4, and 5 indicate haplogroups. Italicized characters with thin curved arrows indicate the geographic place of haplogroup extinction. Thick arrows connecting haplogroups indicate ancestral and derived forms. Thick arrows connecting two different areas indicate the direction of intercontinental migrations. Models I and II propose an Asian origin of YAP+ chromosomes; model III proposes an African origin.

were later confirmed by Hammer et al. (1997).

Recently, one group of researchers analyzed the distribution of 9 (Hammer et al., 1998) and 12 (Karafet et al., 1999) Y-specific biallelic markers in a large number of individuals representing populations from Africa, Asia, Europe, Oceania, and America, and defined 10 (Hammer et al., 1998) and 14 (Karafet et al., 1999) different Y-haplotypes. By using a coalescence analysis, Hammer et al. (1998) confirmed Africa as the continent of origin of modern humans and proposed Asia as the place of origin of the ancestral YAP+ element. On the other hand, these authors estimated a later age for the origin of the ancestral Y-chromosome (150,000–130,000 YBP) and YAP+ chromosomes (60,000–55,000YBP) (see Fig. 7 in Hammer et al., 1998, and Fig. 2 in Karafet et al., 1999). A more recent origin of the *Alu* insert could, at first sight, be in better agreement with the hypothesis of an Asian origin of the YAP+ variant (Altheide and Hammer, 1997; Hammer et al., 1998).

In Karafet et al. (1999), the Native-American-specific C > T transition at the DYS199 locus is proposed to have appeared in the New World approximately 7,600 YBP, a date representing an underestimation, as there is now strong evidence indicating that DYS199*T occurred in Beringia 30,000–20,000 YBP (Underhill et al., 1996; Lell et al., 1997; Bianchi et al., 1998; Santos et al., 1999). Accordingly, it seems valid to evaluate whether the more recent origin proposed by Hammer et al. (1998) and Karafet et al. (1999) for the ancestral YAP+ element also represents an underestimation.

In monogamous populations, the effective population size of autosomal genes is four times greater than those of the haploid Y-

chromosome and mtDNA. Yet, when the mode of hereditary transmission of these two haploid systems is compared, there are clear sex-specific peculiarities that can be only explained by differences between both systems (Seielstad et al., 1998; Pérez-Lezaun et al., 1999). In the case of Y-chromosomes, increased male mortality by violent causes (big game hunting, warfare) (Walker and Lambert, 1989; Seielstad et al., 1998) and the practice of polygyny (Pena et al., 1995; Seielstad et al., 1998) give rise to a lower effective population size than that resulting from the haploid state of the Y-chromosome (Seielstad et al., 1998). Moreover, polygyny and patrilocality (tendency of females to move into the geographic area of their male reproductive partners) (Seielstad et al., 1998; Pérez-Lezaun et al., 1999) do not fit the concept of random mating.

Methods used to estimate the age of coalescence with Y-specific markers usually assume random mating and constant effective population sizes through all the evolutionary time under analysis (Hammer et al., 1998). However, as we discussed above, these two assumptions do not necessarily apply for Y-chromosomes. On the other hand, genetic drift, population subdivision, and full linkage disequilibrium of Y-specific markers give rise to frequent losses of paternal lineages that play an important role in the evolution of Y-chromosomes (Pérez-Lezaun et al., 1999). The coexistence of two biallelic mutations in the branching of the YAP+ haplogroup 3A into 3G (S and PN2, Fig. 1) and in the branching of the haplogroup 4 into 5 (DYS271 and PN1, Fig. 1) gives two examples that very likely illustrate the extinction of paternal lineages (see above).

Due to the reasons detailed above, there is a clear risk of underestimation in the age of origin of ancestral Y-haplotypes. Novel methods of age estimation of Y lineages have been developed elsewhere (Underhill et al., 1996; Bianchi et al 1998). Unfortunately, these methods cannot be used to date YAP+ variants due to the multiple partitioning and monophyletic origins of YAP+ haplogroups, subhaplogroups, and haplotypes. Accordingly, we think it valid to assume that the age of 60,000–55,000 years BP proposed by Hammer et al. (1998) and Karafet et al. (1999) for the origin of YAP+ chromosomes might be an underestimation, and therefore the earlier dating (141,000 years BP) proposed by Hammer (1995) is probably more realistic.

Taking into account the evidence currently available, it is not possible to conclude at present which of the models depicted in Figure 2 is the one best fitting the anthropological genetic data.

## ACKNOWLEDGMENTS

We thank all the colleagues mentioned in Materials and Methods who kindly provided the samples used in this report.

## LITERATURE CITED

Altheide TK, Hammer MF. 1997. Evidence for a possible Asian origin of YAP+ Y chromosomes. Am J Hum Genet 61:462–466.

Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ, Deininger PL. 1994. African origin of human-specific polymorphic *Alu* insertions. Proc Natl Acad Sci USA 91:12288–12292

Bianchi NO, Bailliet G, Bravi CM, Carnese FR, Rothhammer F, Martínez-Marignac VL, Pena SDJ. 1997. Origin of Amerindian Y-chromosomes as inferred by the analysis of six polymorphic markers. Am J Phys Anthropol 102:79–89.

Bianchi NO, Catanesi CI, Bailliet G, Martínez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera RJ, López-Camelo JS. 1998. Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. Am J Hum Genet 63:1862–1871

Bravi CM, Bailliet G, Crisci JV, Bianchi NO. 1997a. Evolution of YAP+ human Y-chromosomes. J Genet Mol Biol 8:65–69.

Bravi CM, Sans M, Bailliet G, Martínez-Marignac VL, Portas M, Barreto I, Bonilla C, Bianchi NO. 1997b. Characterization of mitochondrial DNA and Y-chromosome haplotypes in a Uruguayan population of African ancestry. Hum Biol 69:641–652.

Chakraborty R. 1992. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. Hum Biol 64:141–159.

de Kniff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidon F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Pérez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Roewer L. 1997. Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110:134–140.

Hammer MF. 1994. A recent insertion of an *Alu* element on the Y-chromosome is a useful marker for human population studies. Mol Biol Evol 11:749–761.

Hammer MF. 1995. A recent common ancestry for human Y-chromosomes. Nature 378:376–378.

Hammer MF, Horai S. 1995. Y chromosomal variation and the peopling of Japan. Am J Hum Genet 56:951–962.

Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, Mitchell RJ, Horai S, Jenkins T, Zegura SL. 1997. The geographic distribu-

tion of human Y-chromosome variation. Genetics 145: 787–805.

Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL. 1998. Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol Biol Evol 15:427–441.

Harpending HC, Batzer MA, Gurven M, Jorde LB, Rogers AR, Sherry ST. 1998. Genetic traces of ancient demography. Proc Natl Acad Sci USA 95:1961–1967.

Karafet T, Zegura SL, Vuturo-Brady J, Posukh O, Osipova L, Wiebe V, Romero F, Long JC, Harihara S, Jim F, Dashnyam B, Gerelsaikhan T, Omoto K, Hammer MF. 1997. Y-chromosome markers and trans-Bering Strait dispersals. Am J Phys Anthropol 102:301–314.

Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, Long J, Goldman D, Klitz W, Harihara S, de Knijff P, Wiebe V, Griffiths RC, Templeton AR, Hammer MF. 1999. Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. Am J Hum Genet 64:817–831.

Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Pérez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, de Knijff P, Roewer L. 1997. Ev**Alu**ation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133.

Lell JT, Brown MD, Schurr TG, Sukernik RI, Starikovskaya YB, Torroni A, Moore LG, Troup GM, Wallace DC. 1997. Y-chromosome polymorphisms in Native American and Siberian populations: identification of Native American Y-chromosome haplotypes. Hum Genet 100:536–543.

Pena SDJ, Santos FR, Bianchi NO, Bravi CM, Carnese FR, Rothhammer F, Gerelsaikhan T, Munkhtuja B, Oyunsuren T. 1995. Identification of a major founder Y-chromosome haplotype in Amerindians. Nat Genet 11:15–16.

Pérez-Lezaun A, Calafell F, Comas D, Mateu E, Bosch E, Martínez-Arias R, Clarimón J, Fiori G, Luiselli D, Facchini F, Pettener D, Bertranpetit J. 1999. Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. Am J Hum Genet 65: 208–219.

Santos FR, Bianchi NO, Pena SDJ. 1996. Worldwide distribution of human Y-chromosome haplotypes. Genome Res 6:601–611.

Santos FR, Pandya A, Tyler-Smith C, Pena SDJ, Schanfield M, Leonard WR, Osipova L, Crawford MH, Mitchell RJ. 1999. The Central Siberian origin for native American Y-chromosomes. Am J Hum Genet 64:619–628.

Seielstad M, Herbert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D, Cavalli-Sforza L. 1994. Construction of human Y-chromosomal haplotypes using new polymorphic A to G transition. Hum Mol Genet 3:2159–2161.

Seielstad M, Minch E, Cavalli-Sforza LL. 1998. Genetic evidence for a higher female migration rate in humans. Nat Genet 20:278–280.

Spurdle AB, Hammer MF, Jenkins T. 1994. The Y *Alu* polymorphism in Southern African populations and its relationship to other Y-specific polymorphisms. Am J Hum Genet 54:319–330.

Stoneking M, Fontius JJ, Clifford SL, Soodyall H, Arcot SS, Saha N, Jenkins T, Tahir MA, Deininger PL, Batzer MA. 1997. *Alu* insertion polymorphisms and human evolution: evidence for a larger population size in Africa. Genome Res 7:1061–1071.

Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, Goldstein DB. 1998. Origins of Old Testament priests. Nature 394:138–140.

Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL. 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 93:196–200.

von Haeseler A, Sajantila A, Pääbo S. 1995. The genetical archaeology of the human genome. Nat Genet 14:135–140.

Walker PL, Lambert P. 1989. Skeletal evidence for stress during a period of cultural change in prehistoric California. In: Capasso L, editor. Advances in paleopathology, Journal of Paleopathology: monographic publication 1. Banning, CA: Malki Museum Press. p 207–212.

Whitfield LS, Sulston JE, Goodfellow PN 1995. Sequence variation of the human Y chromosome. Nature 378:379–380.